

匹配人和交通

滴滴出行. 算法工程师. 郭栋

<http://dongguo.me>

Facebook

Netflix

淘宝

百度

匹配

美团

滴滴

匹配问题

	Facebook	Netflix	百度	淘宝	美团	滴滴
匹配对象	人	视频	广告	商品	服务	车(人)
匹配对象的资源限制	无限制	无限制	弱限制	弱限制	弱限制	强限制(独占)

如何匹配？

- 优化目标：转化率
- 策略1：相关性过滤+ 排序(大规模机器学习)？
- 策略2：数学优化问题

匹配人和交通

- 核心目标：最大化乘客的出行体验
- 背景：
 - 出行方式多种多样（专车、出租车、快车（快车拼车）、顺风车、大巴...）
 - 供求常常不均衡，且在时空2个维度差异显著
 - 乘客有自己的偏好
 - 司机是博弈专家

滴滴生态

- 建立一个健康的出行生态
 - 产品导流
 - 动态价格
 - 智能补贴
 - 订单分配

订单如何分配？

- 不同业务线情况不一样
- 出租车
 - 特点：司机的弱控制性
 - 目标：最大化成交率
 - 产品形态：多轮播单，梯度播单距离
 - 策略：整体匹配，组合优化，听抢率预估

定义

$STR_{ij}(t)$: 司机 j 在 t 秒之后抢订单 i 的概率

$STR_{ij}(0)$: 司机 j 抢订单 i 的概率（简化为 STR_{ij} ）

$PSR(i, H_i) = 1 - \prod_{j=1}^{|H_i|} (1 - STR_{d(i)j}(t_{ij}))$: 订单 i 在播单历史 H_i 时被接单的概率

$PSR(i, H_i) \cdot (1 - CAN_i) = (1 - \prod_{j=1}^{|H_i|} (1 - STR_{d(i)j}(t_{ij}))) \cdot (1 - CAN_i)$: 订单 i 在播单历史 H_i 时成交概率

优化目标

	o_1	...	o_j	...	o_N
d_1	x_11	...	x_1j	...	x_1N
...
d_i	x_i1	...	x_ij	...	x_iN
...
d_M	x_M1	...	x_Mj	...	x_MN

$$x_{ij} \in \{0, 1\}, \sum_{j=1}^N x_{ij} = 1, x_{ij} = 1: \text{给司机 } i \text{ 派送订单 } j$$

$$\text{Maximize } \sum_{i=1}^N \{1 - [1 - PSR(i, H(i))] \cdot \prod_{j=1}^M [1 - x_{ij} \cdot STR_{ij}(0)]\} \cdot (1 - CAN_i)$$

$$\text{Giving } x_{ij} \in \{0, 1\}, \sum_{j=1}^N x_{ij} = 1 \quad \forall i \in \{1, 2, \dots, M\}$$

技术总结

- 策略上

- 基于Spark的大数据处理和分布式模型训练自动化流程
- STR预估：十亿级别样本，Logistic Regression + L-BFGS
- 在线学习，深度学习
- 优化求解： $O(m*n)$

- 架构上

- 分布式分单引擎：十万订单十万司机的1s匹配
- 分单全流程的track和监控

其它总结

- 从整体统筹策略设计
- 了解真正的业务目标（知道痛点）
- 算法策略 + 产品策略 + 运营策略
- 简单好过复杂

欢迎加入滴滴

- 和优秀的人一起做有价值的事
- 内推：简历至 guodong@didichuxing.com